

# UNIVERSITY OF BIRMINGHAM

## Research at Birmingham

### Effects of orthographic forms on the acquisition of novel spoken words in a second language

Cerni, Tania; Bassetti, Bene; Masterson, Jackie

DOI:

[10.3389/fcomm.2019.00031](https://doi.org/10.3389/fcomm.2019.00031)

License:

Creative Commons: Attribution (CC BY)

*Document Version*

Publisher's PDF, also known as Version of record

*Citation for published version (Harvard):*

Cerni, T, Bassetti, B & Masterson, J 2019, 'Effects of orthographic forms on the acquisition of novel spoken words in a second language', *Frontiers in Communication*, vol. 4, no. 31.  
<https://doi.org/10.3389/fcomm.2019.00031>

[Link to publication on Research at Birmingham portal](#)

#### **Publisher Rights Statement:**

Cerni T, Bassetti B and Masterson J (2019) Effects of Orthographic Forms on the Acquisition of Novel Spoken Words in a Second Language. *Front. Commun.* 4:31. doi: 10.3389/fcomm.2019.00031

#### **General rights**

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

#### **Take down policy**

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact [UBIRA@lists.bham.ac.uk](mailto:UBIRA@lists.bham.ac.uk) providing details and we will remove access to the work immediately and investigate.



# Effects of Orthographic Forms on the Acquisition of Novel Spoken Words in a Second Language

Tania Cerni<sup>1\*</sup>, Bene Bassetti<sup>2</sup> and Jackie Masterson<sup>3</sup>

<sup>1</sup> Department of Psychology and Cognitive Science, University of Trento, Rovereto, Italy, <sup>2</sup> School of Education, University of Birmingham, Birmingham, United Kingdom, <sup>3</sup> Department of Psychology and Human Development, UCL Institute of Education, London, United Kingdom

## OPEN ACCESS

### Edited by:

Eva Kehayia,  
McGill University, Canada

### Reviewed by:

Marilyn Vihman,  
University of York, United Kingdom  
Caicai Zhang,  
Hong Kong Polytechnic University,  
Hong Kong

### \*Correspondence:

Tania Cerni  
tania.cerni@gmail.com

### Specialty section:

This article was submitted to  
Language Sciences,  
a section of the journal  
Frontiers in Communication

**Received:** 31 October 2018

**Accepted:** 21 June 2019

**Published:** 31 July 2019

### Citation:

Cerni T, Bassetti B and Masterson J  
(2019) Effects of Orthographic Forms  
on the Acquisition of Novel Spoken  
Words in a Second Language.  
Front. Commun. 4:31.  
doi: 10.3389/fcomm.2019.00031

The orthographic forms of words (spellings) can affect word production in speakers of second languages. This study tested whether presenting orthographic forms during L2 word learning can lead speakers to learn non-nativelike phonological forms of L2 words, as reflected in production and metalinguistic awareness. Italian<sub>L1</sub> learners of English as a Second Language (English<sub>L2</sub>) were exposed to English<sub>L2</sub> novel spoken words (pseudowords) and real words in association with pictures either from auditory input only (Phonology group), or from both auditory and orthographic input (Phonology & Orthography group, both groups  $n = 24$ ). Pseudowords and words were designed to obtain 30 semi-minimal pairs, each consisting of a word and a pseudoword that contained the same target consonant, spelled with one letter or with double letters. In Italian double consonant letters represent a long consonant, whereas the English language does not contrast short and long consonants. After the learning phase, participants performed a production task (picture naming), a metalinguistic awareness task (rhyme judgment) and a spelling task. Results showed that the Phonology & Orthography group produced the same consonant as longer in double-letter than in single-letter lexical items, while this was not the case for the Phonology group. The former group also rejected spoken rhymes that contained the same consonant spelled with a single letter in one word and double letters in the other, because they considered these as two different phonological categories. Finally, the Phonology & Orthography group learned more novel words than the Phonology group, showing that orthographic input results in more word learning, in line with previous findings from native speakers.

**Keywords:** word learning, orthographic effects, second language, language production, metalinguistic awareness

## INTRODUCTION

Orthographic input impacts spoken word learning in native, second and novel languages (languages unknown to the participant, including artificial languages), although effects are different. In a native language, those who are exposed to both auditory and orthographic input learn more spoken words than those exposed only to auditory input. This positive effect of orthographic input – sometimes referred to as “orthographic facilitation” (Chambré et al., 2017)—was found in both children (Ehri and Wilce, 1979; Rosenthal and Ehri, 2008; Ricketts et al., 2009) and adults (Nelson et al., 2005; Miles et al., 2016), who memorized the spoken form and meaning of words

more easily if they saw the word's written form. The same was found in child learners of English as a foreign language (Hu, 2008; Vadasy and Sanders, 2015).

Among adult second language learners, orthographic input facilitates the perception of the phonological form of novel L2 words. Thus, Escudero et al. (2008) found that Dutch<sub>L1</sub> speakers of English<sub>L2</sub> could better recognize spoken English<sub>L2</sub> pseudowords containing one of the two vowels /æ/–/ɛ/, a confusable contrast for Dutch<sub>L1</sub> speakers, when in a previous word learning phase the phonological forms of these words were accompanied by their orthographic forms (spellings). However, Escudero et al. (2014) found that the effects of orthographic input depend on the correspondences between the graphemes and phonemes in the native language, such that Dutch<sub>L2</sub> learners were facilitated in Dutch word learning if the word's orthographic form contained grapheme-phoneme correspondences congruent with those in the native language, and inhibited when the two languages' grapheme-phoneme correspondences were incongruent.

In novel languages, the effects of orthographic input on the learning of the phonological forms of words are less clear-cut, ranging from positive to negative to no effects, depending on the difficulty of the sounds to be learned and familiarity of the symbols used to represent them. The first study to investigate the phenomenon (Erdener and Burnham, 2005) found positive effects in English<sub>L1</sub> and Turkish<sub>L1</sub> learners of two novel languages (Irish and Spanish). In particular, orthographic input facilitated the learning of phonological forms in native speakers of the phonologically transparent Turkish language, and learners of the phonologically transparent Spanish as a novel language. While all the languages studied in Erdener and Burnham's (2005) paper were written with the roman alphabet, Showalter et al. (2013) found that even unfamiliar orthographic symbols could improve the perception of spoken words in a novel language. In their study, English native speakers with no experience of Chinese were taught to associate pictures with Chinese pseudowords which varied in lexical tone. Participants who had learned the new spoken words together with their orthographic form in romanised Chinese outperformed those who had only received auditory input in a subsequent word recognition task. However, orthographic input does not seem to help novel language learners when the contrast is difficult. Showalter and Hayes-Harb (2015) carried out a study where English<sub>L1</sub> speakers with no knowledge of the Arabic language learned a set of Arabic pseudoword minimal pairs created to test their perception of the Arabic velar–uvular contrast (e.g., [kubu], [qubu]), which is particularly difficult for English<sub>L1</sub> speakers. Each word was associated with an image, and with either the word's orthographic form (experimental condition) or an unrelated string of Arabic symbols (control condition). Orthographic input did not facilitate the learning of this contrast, even when participants were instructed about the Arabic writing system, or when Arabic symbols were replaced with romanisation. Pytlyk (2011) also found that presenting orthographic input during word learning did not help Chinese phoneme discrimination in English<sub>L1</sub> speakers with no previous exposure to Chinese. There were no differences between presenting Chinese transcriptions

written in the familiar roman alphabet or in an unfamiliar Chinese transcription system. Mathieu (2016) even found a negative effect of orthographic input presentation in English<sub>L1</sub> speakers acquiring a consonantal contrast in Arabic either with the spoken forms alone or else spoken forms plus orthographic forms in one of three scripts—Arabic, cyrillic and a hybrid roman/cyrillic script. Participants who had learned the spoken words with the orthographic forms were less accurate in recovering the target contrast than participants who had only heard auditory input, because familiar scripts had effects when the same grapheme represented a different sound in the native and the novel language. It appears that orthographic input affects the learning of the phonological forms of new words and sounds, with generally positive effects in L2 speakers and more varied effects in learners of novel languages.

Looking at the production of familiar L2 words by experienced L2 speakers, the words' orthographic forms can lead to non-nativelike production (Bassetti, 2007; Hayes-Harb et al., 2010). Effects can include sound addition, as when L2 speakers add an epenthetic sound corresponding to a so-called “silent letter,” such as a [l] in *walk* (Bassetti and Atkinson, 2015). However, the most studied orthographic effects on L2 speech production happen because L2 speakers recode L2 orthographic forms using L1 grapheme-phoneme correspondences, which results in sound substitution, for instance when L2 speakers of American English pronounce a [t] or a [d] in words spelled with letters <t> or <d>, which native speakers produce as flaps (Vokic, 2011). Research by Bassetti and colleagues in particular shows that L2 orthographic forms can lead L2 speakers to establish a phonological contrast in their L2 phonological system that does not exist in the phonological system of the target language. Bassetti (2017) found that Italian English<sub>L2</sub> learners, whose native orthography uses double consonant letters to represent geminates (long consonants), tend to produce the same English<sub>L2</sub> consonant as longer when spelled with double consonants than with singleton consonants, for instance producing a longer /t/ in *kitty* than in *city*. This was found in a reading aloud task but also in a delayed word repetition task with or without orthographic input presented before the repetition. The fact that the effect appeared also when the spelling of the word was not available could attest to an orthography-influenced L2 phonological representation, or to the activation of the L2 orthographic representation during L2 speech production. In support of this interpretation, Italian speakers of English<sub>L2</sub> were found to produce minimal pairs distinguished by a geminate or singleton consonant such as /'fin:ɪʃ/—/'finɪʃ/ (*Finnish-finish*, both /'finɪʃ/ in British English; Bassetti et al., 2018). Furthermore, they rejected rhymes containing the same consonant written with a singleton letter or with double letters (e.g., *very-merry*) because they considered these two phonological categories, a singleton and a geminate consonant, as in Italian<sub>L1</sub> (Bassetti et al., under revision).

Previous research has found, then, that orthographic forms can affect the perception and production of L2 words. However, these effects could be due to repeated exposures rather than being established at the point of first learning words. In the present study we aimed to investigate whether orthographic

effects are found in the very early stages of acquiring vocabulary items. We investigated whether Italian<sub>L1</sub> speakers of English<sub>L2</sub> would learn a spoken novel word in English as containing a geminate (long) consonant if the word was presented with an orthographic form that contains double consonant letters. This outcome could come about as the result of recoding the English word's orthographic form according to Italian<sub>L1</sub>'s grapheme-phoneme correspondence rules. The recoded orthographic form would then interact with the phonological form from the L2 spoken input—which contains a short consonant—resulting in a conflict between the two forms—and possibly a perceptual illusion—and result in a phonological representation containing a long consonant.

To test this, we compared the learning of pseudowords in two groups of Italian English<sub>L2</sub> learners. One group received only phonological input during the learning phase (Phonology group) and the other received orthographic input simultaneously with the phonological input (Phonology & Orthography group). We then assessed participants' production, metalinguistic awareness and spelling of the target sounds. Items for the learning task comprised word and pseudoword pairs. The pseudoword in each pair was created by replacing the onset (initial consonant or consonant cluster) of the word (e.g., /'prɪnɪʃ/-/'fɪnɪʃ/, *prinish-finish*).

The first aim of the study was to assess effects of orthography on the spoken production of novel words. To address this aim, participants learned spoken stimuli in association with pictures. We then used a picture naming task to compare the target consonant duration ratio in the Phonology and Phonology & Orthography groups. Target consonant duration ratio was obtained by dividing the duration of the target consonant in the first word of the pair by the duration of same target consonant in the other word, a procedure that has been used in previous studies (Bassetti et al., 2018; Bassetti et al., under revision). There were three types of pairs: CC<sub>pw</sub>-C<sub>w</sub> pairs consisted of a pseudoword where the target consonant was spelled with double letters and a word where it was spelled with a single letter, such as *prinnish-finish*; CC<sub>w</sub>-C<sub>pw</sub> pairs consisted of a word where the target consonant was spelled with double letters and a pseudoword where it was spelled with a single consonant letter, such as *Finnish-prinish*; C<sub>pw</sub>-C<sub>w</sub> pairs consisted of a pseudoword and a word both containing the target consonant spelled with a single letter (e.g., *prinish-finish*). We had different predictions for each group for each type of pair, as follows. With regards to the Phonology & Orthography group, we predicted that they would have high ratios in pairs where pseudowords were spelled with double consonants and words with a single consonant (CC<sub>pw</sub>-C<sub>w</sub> pairs, e.g., *prinnish-finish*), because they were expected to produce responses where the consonant in the pseudoword was longer than the consonant in matched words. For this group, we predicted high ratios also in pairs consisting of a word spelled with double consonants and a pseudoword spelled with a single consonant (CC<sub>w</sub>-C<sub>pw</sub> pairs, e.g., *Finnish-prinish*), because we predicted that they would produce responses where the consonant in the word was longer than the consonant in the matched pseudoword. Finally, we predicted a ratio of around one

for pairs where both pseudoword and word were spelled with a single consonant letter (C<sub>pw</sub>-C<sub>w</sub> pairs, e.g., *prinish-finish*), as both consonants would be produced with similar lengths.

A related aim was to assess how Italians would categorize and learn consonants in pseudowords in the absence of orthography. We had two different sets of predictions for the speech production of the Phonology group. If, in the absence of orthographic input, Italians categorize all English consonants as singletons, we predicted a ratio of around one for CC<sub>pw</sub>-C<sub>w</sub> and C<sub>pw</sub>-C<sub>w</sub> pairs, as both pseudowords and words spelled with a single letter would be produced as short, and we predicted a high ratio for CC<sub>w</sub>-C<sub>pw</sub> pairs, as the word spelled with double consonants would be produced as longer than the pseudoword. On the other hand, Bassetti (2017) argued that Italians may perceive the duration of English consonants as being on the boundary between short and long consonants, and perceive and produce such consonants as long or short depending on their spelling. If this hypothesis is correct, then in the absence of orthographic information Italians should categorize English consonants as either singleton or geminate. In line with this hypothesis, the Phonology group should produce all pairs with ratios higher than one because all pseudowords would be variably produced as either long or short consonants. The Phonology group would therefore have a lower ratio than the Phonology & Orthography group in CC<sub>pw</sub>-C<sub>w</sub> pairs, as the Orthography group would produce these consistently with a long consonant in the pseudoword and a short consonant in the word, whereas the Phonology group would produce some pseudowords with a long consonant and some with a short consonant. With C<sub>pw</sub>-C<sub>w</sub> pairs, the Phonology group would have a higher ratio than the Phonology & Orthography group's ratio of one. Finally, looking at CC<sub>w</sub>-C<sub>pw</sub> pairs, the Phonology group would have a lower ratio than the Phonology & Orthography group, because the latter would produce all pseudowords with short consonants, but the former would produce some with a long consonant and some with a short one.

The second aim of the study was to test whether exposure to orthographic input at the point of initial word learning affects metalinguistic awareness. Fewer studies had demonstrated that the presence of the orthographic form during L2 learning affects phonological awareness in comparison to the auditory input alone. For instance, Detey and Nespoulous (2008) found that L2 written forms led Japanese<sub>L1</sub> speakers during syllabic segmentation to add epenthetic vowels in French<sub>L2</sub> pseudowords containing consonant clusters that are not legal in L1 Japanese. This error did not occur if the pseudowords were learned without the written form (see also Young-Scholten et al., 1999). In the present study, we used a rhyme judgment task, in which participants had to decide whether a pair of lexical stimuli constituted a rhyme. The task had been previously used by Bassetti et al. (under revision) to test the effects of number of consonant letters using pairs of spoken words, and in the present study we included pseudowords. We predicted that the Phonology & Orthography group would incorrectly reject a pseudoword-word pair such as *prinnish-finish* as rhymes, because they would consider the consonant in the pseudoword as a geminate.



As the final aim, we tested whether exposure to orthographic forms increases the number of spoken words learned, compared with auditory input only. There is evidence that learning spoken words together with orthographic forms results in more word learning in a native language (Nelson et al., 2005; Miles et al., 2016). However, the evidence of this facilitative effect in a second language is limited to child beginner learners of English (Hu, 2008; Vadasy and Sanders, 2015). We predicted that the Phonology & Orthography group would learn more spoken novel words, compared to the Phonology group, and therefore show more accurate performance in the picture naming task.

## MATERIALS AND METHODS

### Participants

Forty-eight Italian high-school learners of English were randomly assigned to one of two groups: a Phonology group and a Phonology & Orthography group (both  $n = 24$ ; two additional participants were eliminated because they failed to complete the learning phase described below). Both groups learned the same English spoken words and pseudowords from auditory input during the experiment, but the Phonology & Orthography group also received orthographic input.

Participants were native speakers of the Roman variety of Standard Italian, who were attending the third ( $n = 41$ ) or fourth ( $n = 7$ , of which 4 in the Phonology group) year in one of two state-run high schools in Rome, a classical and a scientific high school. None of the participants knew another language with contrastive consonant length. One student reported being dyslexic, but performed similarly to the others. Participation in the study was voluntary and was rewarded with book vouchers.

Participants were studying English as a compulsory school subject for 3 h a week, using British English textbooks. They had been studying English at school for 10.2 years ( $SD = 1.3$ ; Phonology group:  $M = 10.4$ ,  $SD = 1.4$ ; Phonology & Orthography group:  $M = 10.0$ ,  $SD = 1.2$ ). About half (56%) of participants had studied English in extra-scholastic settings with native teachers (Phonology group:  $n = 15$ ;  $Med = 10$  months, range = 4–64; Phonology & Orthography group:  $n = 12$ ;  $Med = 16.5$  months, range = 1–72. Kruskal-Wallis test:  $\chi^2(1) = 0.54$ ,  $p = 0.46$ ). About half (54%) had never been in an English-speaking country, the others reported a few weeks of study abroad (Phonology group:  $n = 11$ ,  $Med = 2.7$  weeks, range = 2–10; Phonology & Orthography group:  $n = 11$ ,  $Med = 2$  weeks, range = 2–4. Kruskal-Wallis test:  $\chi^2(1) = 0.01$ ,  $p = 0.90$ ).

Participants reported spending much more time listening to English than reading it (Phonology group: Listening— $Med = 3$  h per week, range = 0–28; Reading— $Med = 1$ , range = 0–10. Phonology & Orthography group: Listening— $Med = 4.5$ , range = 0–30; Reading— $Med = 1$ , range = 0–30). They spoke English for < 1 h a week (Phonology group:  $Med = 0$ , range = 0–10; Phonology & Orthography group: Reading— $Med = 0.7$ , range = 0–6). All but one of the participants considered a native-like English pronunciation important, very important, or extremely important. Overall, they preferred a British English accent to an American English one.

### Stimuli

Materials consisted of 20 English words and 20 pseudowords (see **Supplementary Table 1**). All lexical items were monomorphemic and disyllabic, and contained a target consonant—[p], [t], or [n]—in post-tonic intervocalic position (see Bassetti, 2017; Bassetti et al., under revision). Real words were frequent lexical items for participants, as confirmed by their teachers. Pseudowords were created by changing the initial consonants of their paired word, for instance creating *prinnish* from *Finnish*. For each pseudoword C- and CC- versions were created, for instance *prinnish* and *prinish*. Half of the items were spelled with double consonants (CC) and half with a single consonant (C). There were therefore ten each of CC-words, CC-pseudowords, C-words, and C-pseudowords. Lexical items were nouns for tools, animals or plants (real words also included three adjectives), and were associated with an image of their referent.

These 40 words were used to create 30 semi-minimal pairs, each consisting of a word and a pseudoword that contained the same target consonant in the same VCV rhyme. Target consonant spelling was manipulated to obtain three types of word-pseudoword pair as follows. In  $CC_{pw}-C_w$  pairs, the target consonant was spelled with double letters in the pseudoword and a single consonant letter in the word, e.g., /'prɪnɪʃ/-/'fɪnɪʃ/ (*prinnish-finish*). In  $C_{pw}-C_w$  pairs, the target consonant was spelled with a single consonant letter in both word and pseudoword, e.g., /'prɪnɪʃ/-/'fɪnɪʃ/ (*prinish-finish*). In  $CC_w-C_{pw}$  pairs, the target consonant was spelled with double letters in the word and a single consonant letter in the pseudoword, e.g., /'fɪnɪʃ/-/'prɪnɪʃ/ (*Finnish-prinish*). Creating pseudoword-word pairs was more efficient than creating pairs of pseudowords, because it allowed us to test our hypothesis with half as many pseudowords as would be needed to use pseudoword-pseudoword pairs. This was important, considering that learning new words in this type of experimental setting is very time-consuming (Escudero et al., 2008). By using the same pseudowords in  $CC_w-C_{pw}$  and  $C_{pw}-C_w$  pairs, it was possible to use the 40 lexical items to obtain ten each of  $CC_{pw}-C_w$ ,  $CC_w-C_{pw}$ , and  $C_{pw}-C_w$  pairs.

To reduce the risk of fatigue and frustration, each participant only learned 10 pseudowords (five C- and five CC-pseudowords). They also learned the association between a picture and 15 words (ten C- and five CC-words). To achieve this, each participant learned one of two lists, each one containing five each of  $CC_{pw}-C_w$ ,  $CC_w-C_{pw}$  pairs, and  $C_{pw}-C_w$  pairs. Each list contained only the C- or the CC- version of each pseudoword. For instance, *prinnish* appeared in the first list within the  $CC_{pw}-C_w$  pair *prinnish-finish*, and *prinish* in the second list within the  $C_{pw}-C_w$  pair *prinish-finish*. The list containing the C-pseudoword also contained the corresponding CC word within a CC-C pair (for instance, the list with *prinish-finish* also contained *Finnish-prinish*). The two lists were counterbalanced between participants and groups.

Lexical items were recorded by a female Southern British English native speaker and English language teacher, who read aloud words and pseudowords spelled with a singleton consonant. She was instructed to speak clearly and slightly slowly, and to pronounce pseudoword rhymes as in the corresponding

word. Recording took place in a sound-attenuated room using a Rode NT2-A microphone connected to an Alesis Multimix 12 Firewire mixer.

Each lexical item was paired with a color line drawing (see **Figure 1**). Pictures were used both to illustrate the meaning of the lexical item during the learning phase, and to allow us to elicit spoken production in the testing phase (see Showalter and Hayes-Harb, 2015). Words were paired with images of their referent, and pseudowords with images of unusual tools, plants or animals, whose nouns were highly infrequent in both British English and Italian. The C- and CC-versions of the same pseudoword were paired with the same image. Images were selected from the Art Explosion library (Nova Development, 2004).

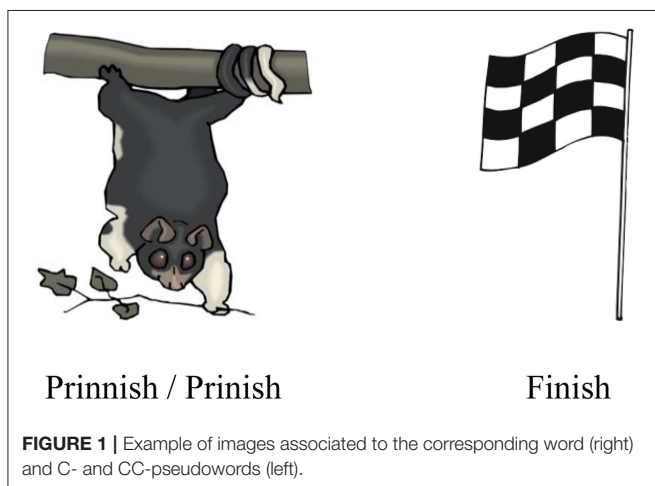
## Tasks and Procedure

Participants first took part in a learning session, which consisted of a learning phase and test. Following this, there were three tasks: picture naming, rhyme judgement and spelling, administered in fixed order. Participants were tested individually in a quiet room at their school, with the researcher (the first author) present to assist, and they could pause whenever they wished during and between tasks. The whole session (including learning and experimental sessions) took ~ 1 h.

Participants were told that they would learn new words, but they were not told that the stimuli included pseudowords. The researcher notified participants of the presence of the pseudowords at the end of the experimental session once all the tasks had been completed, and gave them a list of the images used during the tasks and the associated printed real words.

### Learning Phase

In the learning phase, participants learned the 25 lexical items by seeing an image of the referent and hearing the lexical item. Participants were told that they would be asked first to memorize words paired with images, and that some of the words would be unfamiliar to them. They were instructed that following the memorization task they would be asked to recall the verbal labels when the images were presented on their own. The Phonology & Orthography group also saw the lexical item's written form simultaneously with the auditory input.



A trial started with the presentation of a fixation point in the center of the computer screen, followed after 500 ms by an image, which remained on the screen for 3,500 ms. 1,500 ms after the onset of the image, the audio recording was played over headphones. The Phonology & Orthography group also saw the orthographic form, which appeared under the image. The onset of the orthographic form coincided with the presentation of the auditory word (i.e., 1,500 ms after the onset of the image). There was an interval of 500 ms between trials.

There were four blocks of trials; in each block stimuli were presented twice, making a total of 50 trials per block. In this way each word was presented a total of eight times across the learning phase. Trials were presented in random order within the blocks.

### Learning Test

The learning test was used both to test whether the participant had learned the lexical items, and to revise unlearned items. A trial started with a fixation point in the center of the computer screen, followed after 500 ms by an image. The participant pronounced the noun aloud into a microphone. In order to check the accuracy of their response, the participant pressed the space bar. The image appeared again, and 1,000 ms after the image onset the recording was played over the headphones. The Phonology & Orthography group also saw the orthographic form, which was presented underneath the image. The image remained on the screen for a total of 2,000 ms, then the question “Did you remember the word correctly?” appeared on the screen, and participants answered by clicking on a tick (on the right) or a cross (on the left), both of which were presented under the written question. These answers were collected as a self-reported measure of accuracy. Each lexical item appeared four times, once in each of four blocks of trials, in random order. After each block, the screen displayed the percentage of correct answers for that block. All participants saw the four blocks, regardless of the percentage of correct answers.

### Picture Naming Task

On each trial, a fixation point appeared in the center of the screen and was replaced after 500 ms by an image. The participant produced the corresponding noun aloud for three repetitions in the carrier phrase “The word \_\_\_ should follow,” and clicked on an on-screen button after each repetition. The carrier phrase was used to keep the target word in the nuclear position within the intonational unit (Bassetti, 2017). The three repetitions were used to calculate a mean duration in order to increase reliability (Flege, 1995; Bassetti, 2017). The button presses were used to ensure that participants repeated each item three times (Bassetti, 2017). The utterances were recorded for later analysis. Each image was presented once and in random order, for a total of 25 trials.

### Rhyme Judgment Task

In order to test whether participants considered the consonants spelled with double consonants as geminates, they were asked in the rhyme judgement task to judge whether two lexical items rhymed. CC<sub>pw</sub>-C<sub>w</sub> rhymes consisted of a pseudoword-word near-minimal pair, where both items had the same VCV rhyme, which was spelled with double consonants in the pseudoword

and with a singleton consonant in the word, for instance /'prɪnɪʃ/-/fɪnɪʃ/ (*prinnish-finish*). On each trial participants saw two images side by side in the upper part of the computer screen. They were asked to recall the corresponding nouns, and decide whether they rhymed by clicking on one of two buttons below the image, a tick on the right for a rhyme, or a cross on the left for a non-rhyme. Images appeared on the left or the right interchangeably (half of the time in one of the two positions). A 500 ms black screen was used as a pause between trials. Each participant saw 20 trials: five CC<sub>pw</sub>-C<sub>w</sub> rhymes (CC pseudoword-C word), five C<sub>pw</sub>-C<sub>w</sub> rhymes (C pseudoword-C word), five CC<sub>w</sub>-C<sub>pw</sub> rhymes (CC word-C pseudoword) and five non-rhyme fillers. The fillers comprised two real words, both spelled with singleton consonant (e.g., *mini-many*). Fillers were used to add non-rhyming pairs, and to reduce the number of CC-C pairs, thereby reducing the risk of participants guessing the aim of the experiment.

### Spelling Task

A spelling task was used to test whether the participants in the Phonology & Orthography group correctly remembered the spelling of the 25 lexical items, and to assess how the participants in the Phonology group spelled them. After a 500 ms fixation point, participants saw an image in the center of the screen, and typed its noun on a fixed-length line below the image. The response replaced the line. The participant pressed the return key to start the next trial. Each of the 25 images appeared once, in random order. There were no time limits, and participants were allowed to delete and retype the responses up to the beginning of the next trial. If they could not remember the word, they were allowed to skip the trial, but they were encouraged to always try to type an answer.

### Equipment

All tasks were run in OpenSesame 3.1.9 Jazzy James (Mathôt et al., 2012), which managed randomization and recorded keyboard and mouse responses. Auditory input was presented over an AKG HSD171 headset. Participants' productions were recorded using a Zoom H4N Pro digital recorder connected to the headset's dynamic microphone.

### Analysis

#### Acoustic Analysis

In order to analyze the picture naming task data, a trained phonetician measured the duration of each target sound using Praat software (Boersma and Weenink, 2016) following the standard procedure described in Bassetti et al. (2018).

#### Statistical Analysis

Data were analyzed using R software, version 3.4.4 (RStudio Team, 2018) with RStudio 1.1.447 (R Core Team, 2018). All the following analyses were performed using (linear or generalized) mixed effect models, with package *lmerTest* (Kuznetsova et al., 2017—*lme4* version: Bates et al., 2015b). As a first step, all the models included all the fixed effects of interest. Then, model reduction was performed through likelihood ratio test (Baayen et al., 2008). Initially, the maximal random effects

structure was considered (Barr et al., 2013). In case of failure of convergence or overfitting (random effects were perfectly collinear), we proceeded with model reduction following Bates et al. (2015a). In the Results section we report only the final fixed and random effect structures. *P*-values for *t*-statistics were obtained using Satterthwaite's method for denominator degrees of freedom (provided by the *lmerTest* package). Conditional and marginal R<sup>2</sup> were calculated using function *r.squaredGLMM* in the *MuMIn* package (Bartoń, 2018), whereas Tukey's *post-hoc* contrasts were performed using function *contrast* in the *lsmeans* package (Lenth, 2016).

#### Learning test

A generalized linear mixed model with binomial error distribution was used to analyze the number of spoken pseudowords learned in the learning test. The dependent variable was whether the participant rated their answer as correct (coded as 1) or incorrect (coded as 0) after hearing the correct answer. The responses of five Phonology group participants were lost due to technical issues.

#### Picture naming task

4.9% of the data (59 responses out of a total of 1,200) were lost because the answer was missing or mispronounced, or the recording was not suitable for acoustic analysis due to background noise or interruptions. An additional 4.8% (58 responses) were removed because the response for the corresponding item in the Spelling Task was incorrect (misspelled real words, pseudowords spelled with wrong intervocalic context of the target CC-C). This removal was considered as necessary because orthographic effects on consonant production were only expected when the speaker knew the correct correspondence between the pictures and the items and the correct letters in the items (i.e., target consonants and intervocalic context, see *Spelling Task* paragraph below in this section for a more detailed rationale of error categorization in pseudowords). The final dataset contained 1,083 items. A mean duration of the target sound was calculated on the three repetitions made by the participants for each item. Data analysis was performed on consonant duration ratios which were calculated for each participant for each CC<sub>pw</sub>-C<sub>w</sub>, C<sub>pw</sub>-C<sub>w</sub>, and CC<sub>w</sub>-C<sub>pw</sub> pair. For instance, for the CC<sub>pw</sub>-C<sub>w</sub> pairs *prinnish-finish*, the ratio was calculated by dividing the duration of [n] in *prinnish* by the duration of [n] in *finish*.

Outliers were considered as ratios that were beyond the 99th percentile and below the 1st percentile (2.05% of data). The remaining ratios were log-transformed to approximate normal distribution and examined with linear mixed regression models. Model details are provided in the Results section.

#### Rhyme judgment task

Pairs that contained words and pseudowords which were misspelled in the Spelling Task were removed from the analysis (15.5% of data, see Picture Naming Task above and Spelling Task below for the rationale). After removing incorrect spellings and discarding the fillers pairs, 608 pairs (out of a total of 720) were included in the analysis. We coded correct answers as 1 and

incorrect answers as 0. For the analysis we used a generalized mixed model with binomial error distribution.

### Spelling task

A generalized linear mixed model with binomial error distribution was used to analyze the number of pseudowords spelled correctly in the Spelling Task. Responses were coded as 1 if correct, or 0 if incorrect, and these values were used as the dependent variable in the model. We coded pseudoword spellings as incorrect if the participant had not provided an answer, or answered with a lexical item other than the one represented by the picture, or provided a spelling that was substantially different from the phonological form of the target, for instance if the target consonant was surrounded by a wrong vocalic context. With regards to the spelling of the target consonant, we treated the two participant groups differently. If the target consonant was spelled with the incorrect number of letters, we coded the answer as incorrect in the Phonology & Orthography group. For the Phonology group, we accepted both single and double consonants as correct, because the group had not learned the correct spelling and both spellings are acceptable.

## RESULTS

### Learning Test

The first task participants performed measured the number of pseudowords they learned. **Figure 2** shows the mean proportion of pseudowords participants in the two groups reported having correctly produced over the four learning blocks. In the generalized linear mixed model, we inserted as fixed effects

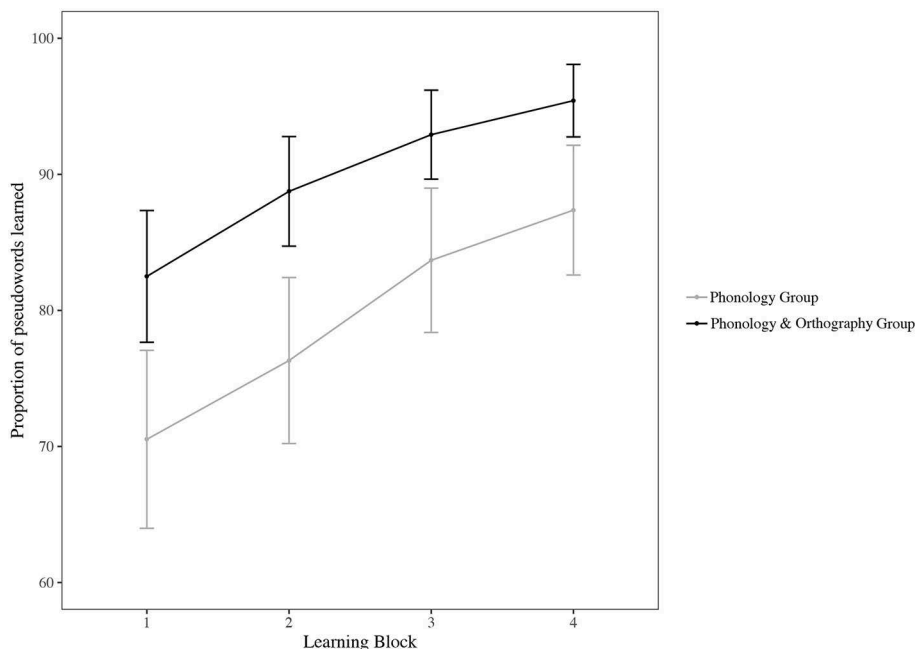
Group (Phonology, Phonology & Orthography) and Block (1, 2, 3, and 4) and their interaction. As random effects we used the intercepts for subjects and items.

The interaction was not significant and was removed from the model, so the final model included the effects of Block and Group. As **Table 1** shows, the Phonology & Orthography group learned more pseudowords than the Phonology group overall, and the lack of a significant interaction shows that this effect applied across the four learning blocks. Both groups learned increasingly more pseudowords over the four learning blocks.

**TABLE 1 |** Results of mixed-model analysis of the effects of group (Phonology, Phonology & Orthography) and learning block (1, 2, 3, 4) on number of pseudowords learned.

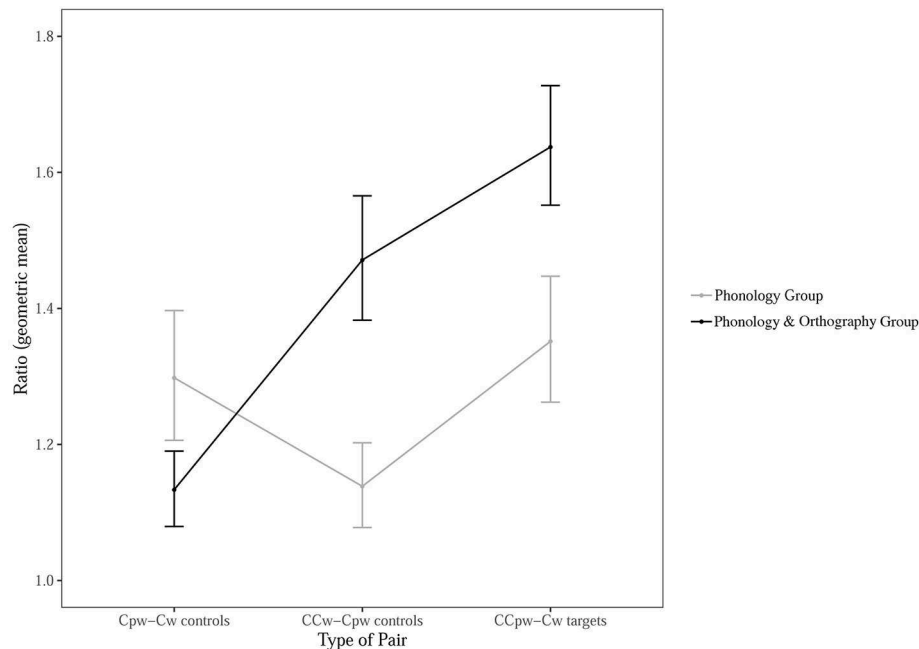
Random effects		Variance	SD		
Participants	Intercept	2.42	1.55		
Word	Intercept	0.47	0.69		
Fixed effects		Estimate	SE	z-value	p
Intercept		1.21	0.42	2.85	0.004**
Group (P&O)		1.21	0.52	2.31	0.020*
Block					
	(Block 2)	0.53	0.20	2.68	0.007**
	(Block 3)	1.15	0.22	5.29	<0.001***
	(Block 4)	1.59	0.24	6.70	<0.001***
Marginal R <sup>2</sup> :		0.10	Conditional R <sup>2</sup> :	0.52	

\* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ .



**FIGURE 2 |** Mean percentage of pseudowords recalled and correctly produced as a function of group (Phonology, Phonology & Orthography) and learning block (1, 2, 3, 4). Bars represent 95% CIs.





**FIGURE 3 |** The geometric mean of consonant duration ratio as a function of group (Phonology, Phonology & Orthography) and type of pair ( $C_{pw}-C_w$ ,  $CC_w-C_{pw}$ ,  $CC_{pw}-C_w$ ). Bars represent 95% CIs.

## Picture Naming Task

This was the critical task because it tested whether orthographic input affected the phonological form of the item that was learned. **Figure 3** shows the results. In line with predictions, for the Phonology & Orthography group, the geometric mean of consonant duration ratios in  $CC_{pw}-C_w$  pairs was 1.64 ( $CI [1.55-1.73]$ ), while the mean ratio of the  $C_{pw}-C_w$  pairs was close to one ( $M = 1.13$ ,  $CI [1.08-1.19]$ ), and the mean ratio of the  $CC_w-C_{pw}$  pairs was close to the mean ratio of the  $CC_{pw}-C_w$  pairs ( $M = 1.47$ ,  $CI [1.38-1.57]$ ). For the Phonology group the geometric mean of the consonant duration ratios was 1.35 in  $CC_{pw}-C_w$  pairs ( $CI [1.26-1.45]$ ), 1.30 in  $C_{pw}-C_w$  pairs ( $CI [1.21-1.40]$ ), and 1.14 in  $CC_w-C_{pw}$  pairs ( $CI [1.08-1.20]$ ).

The linear mixed model included as fixed effects Group (Phonology, Phonology & Orthography), Type of Pair ( $CC_{pw}-C_w$ ,  $C_{pw}-C_w$  vs.  $CC_w-C_{pw}$  pair) and their interaction. As random effect we used the intercepts for subjects and items and the by-subject random slope for Type of Pair. The results are summarized in **Table 2**.

In confirmation of our hypothesis, multiple contrasts within groups with Tukey adjustment revealed that in the Phonology & Orthography group  $CC_{pw}-C_w$  pairs had a higher ratio than  $C_{pw}-C_w$  pairs ( $t = 5.54$ ,  $p < 0.001$ ) and a similar ratio as  $CC_w-C_{pw}$  pairs ( $p > 0.05$ ).  $C_{pw}-C_w$  and  $CC_w-C_{pw}$  pairs differed significantly ( $t = -3.37$ ,  $p = 0.004$ ). These results revealed that the participants in the Phonology & Orthography group produced longer consonant durations for CC words and pseudowords than the consonant durations for the paired C words and pseudowords. For the Phonology group there were no

**TABLE 2 |** Results of mixed-model analysis of the effects of group (Phonology, Phonology & Orthography) and type of pair ( $CC_{pw}-C_w$ ,  $C_{pw}-C_w$ ,  $CC_w-C_{pw}$ ) on consonant duration ratio.

Random effects		Variance	SD	Corr	
Participants	Intercept	0.01	1.10		
	Type of pair ( $C_{pw}-w$ )	0.02	0.13	-0.31	
	( $CC_w-C_{pw}$ )	0.01	0.10	-0.27	-0.83
Word	Intercept	0.01	0.11		
Residual		0.06	0.25		
Fixed effects		Estimate	SE	t-value	p
Intercept		0.28	0.05	5.87	<0.001***
Group (P&O)		0.21	0.05	4.43	<0.001***
TRIAL CONDITION					
( $C_{pw}-C_w$ )		-0.39	0.07	-0.57	ns
( $CC_w-C_{pw}$ )		-0.15	0.07	-2.32	0.025*
GROUP*TRIAL CONDITION					
P&O* $C_{pw}-C_w$		-0.32	0.06	-4.94	<0.001***
P&O* $CC_w-C_{pw}$		0.04	0.06	0.75	ns
Marginal $R^2$ : 0.17		Conditional $R^2$ :		0.42	

\* $p < 0.05$ ; \*\*\* $p < 0.001$ .

statistical differences for  $CC_{pw}-C_w$  and  $C_{pw}-C_w$ ,  $CC_{pw}-C_w$  and  $CC_w-C_{pw}$ ,  $C_{pw}-C_w$  and  $CC_w-C_{pw}$  pairs.

Multiple comparisons between groups showed that the Phonology & Orthography group produced a longer ratio for

both  $CC_{pw}-C_w$  ( $t = 4.43$ ,  $p < 0.001$ ) and  $CC_w-C_{pw}$  ( $t = 4.91$ ,  $p < 0.001$ ) compared with the Phonology group.  $C_{pw}-C_w$  pair ratios were lower for the Phonology & Orthography group than the Phonology group ( $t = -2.04$ ,  $p = 0.049$ ).

## Rhyme Judgment Task

In the Phonology & Orthography group the mean percentage of correct responses was 69% for  $CC_{pw}-C_w$  pairs ( $SD = 46$ ,  $CI = 8$ ), 93% for  $C_{pw}-C_w$  pairs ( $SD = 26$ ,  $CI = 5$ ), and 64% for  $CC_w-C_{pw}$  pairs ( $SD = 48$ ,  $CI = 9$ ). In the Phonology group the mean percentage was 69% for  $CC_{pw}-C_w$  rhymes ( $SD = 47$ ,  $CI = 9$ ), 77% for  $C_{pw}-C_w$  rhymes ( $SD = 42$ ,  $CI = 9$ ) and 87% for  $CC_w-C_{pw}$  rhymes ( $SD = 0.33$ ,  $CI = 0.7$ ). **Figure 4** summarizes the results.

In the generalized linear mixed model, we inserted as fixed effects Group (Phonology, Phonology & Orthography) and Type of Pair ( $CC_{pw}-C_w$ ,  $C_{pw}-C_w$ ,  $CC_w-C_{pw}$  pair) and their interaction. As random effects we used the intercept for subjects and items (words and pseudowords) and the by-subject random slope for Type of Pair. The results are summarized in **Table 3**.

Multiple contrasts within groups with Tukey adjustment revealed that the Phonology & Orthography group was more accurate with  $C_{pw}-C_w$  than  $CC_{pw}-C_w$  rhymes ( $z = 3.42$ ,  $p = 0.002$ ), and  $CC_w-C_{pw}$  accuracy did not differ from  $CC_{pw}-C_w$  accuracy ( $p > 0.05$ ).  $C_{pw}-C_w$  and  $CC_w-C_{pw}$  rhymes differed significantly ( $z = -3.98$ ,  $p < 0.001$ ). Participants in the Phonology & Orthography group erroneously rejected CC-C rhymes (both  $CC_{pw}-C_w$  and  $CC_w-C_{pw}$ ) more often than C-C rhymes, presumably because they considered consonants spelt with a single letter and those spelt with double letters as different phonemes.

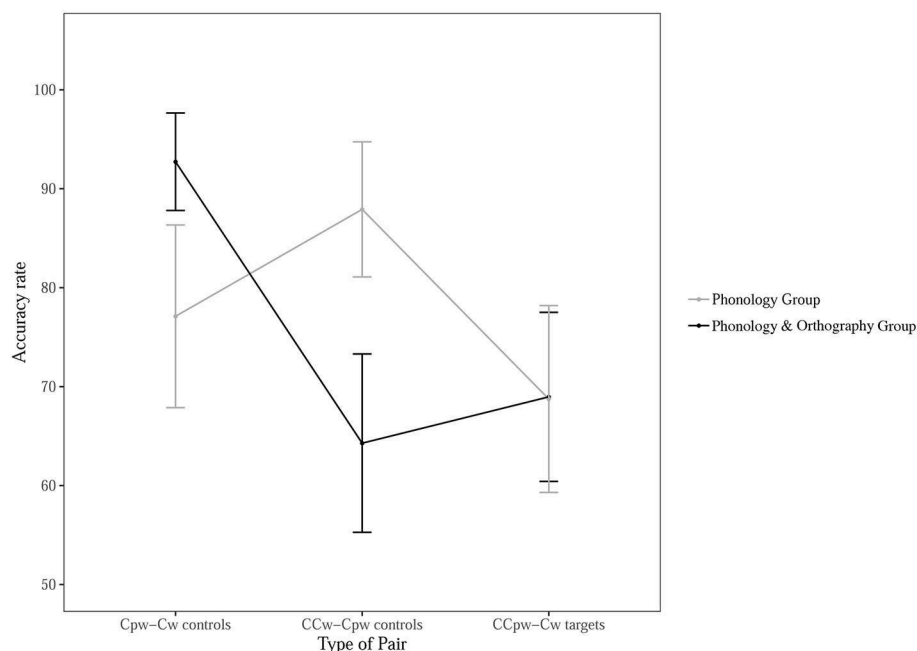
In the Phonology group accuracy was lower with  $CC_{pw}-C_w$  than  $CC_w-C_{pw}$  pairs ( $z = -2.44$ ,  $p = 0.04$ ). There were no differences in accuracy between  $CC_{pw}-C_w$  and  $C_{pw}-C_w$  pairs and between  $CC_w-C_{pw}$  and  $C_{pw}-C_w$  pairs ( $p > 0.05$ ).

Multiple contrasts between groups showed that accuracy did not differ between groups for  $CC_{pw}-C_w$  pairs ( $p > 0.05$ ), but the Phonology & Orthography group was more accurate with  $C_{pw}-C_w$  pairs ( $z = 2.79$ ,  $p = 0.005$ ) and less accurate with  $CC_w-C_{pw}$  pairs ( $z = -3.33$ ,  $p < 0.001$ ) compared with the Phonology group.

## Spelling Task

In the spelling task, the average number of pseudowords spelled correctly was higher in the Phonology & Orthography group ( $M = 97\%$ ,  $SD = 18$ ) than in the Phonology group ( $M = 86\%$ ,  $SD = 35$ ). In the generalized linear mixed model we inserted Group (Phonology, Phonology & Orthography) as fixed effect, and the intercept for participants and pseudowords as random effects. The final model, shown in **Table 4**, reveals that the Phonology & Orthography group spelled more pseudowords correctly than the Phonology group.

In order to understand whether participants perceive English consonants as short or long in words whose spelling they do not know, we analyzed the correct pseudoword spellings produced by the Phonology group. The Phonology group spelled just over half of pseudowords with double consonants (59%, or 122 out of 206 valid pseudoword spellings). A Poisson generalized linear model on count data as the dependent variable and Spelled Consonant (CC vs. C) as fixed effect revealed that the Phonology group spelled pseudowords with double consonants



**FIGURE 4 |** The proportion of correct responses as a function of group (Phonology, Phonology & Orthography) and type of pair ( $C_{pw}-C_w$ ,  $CC_w-C_{pw}$ ,  $CC_{pw}-C_w$ ) in the rhyme judgement task. Bars represent 95% CIs.

**TABLE 3 |** Results of mixed-model analysis of the effects of group (Phonology, Phonology & Orthography) and type of pair ( $CC_{pw}-C_w$ ,  $C_{pw}-C_w$ ,  $CC_w-C_{pw}$ ) on mean percentage of correct responses in the rhyme judgment task.

Random effects		Variance	SD	Corr	
Participants	Intercept	0.69	0.83		
	Type of pair				
	( $C_{pw}-C_w$ )	1.24	0.11	−0.80	
Word	( $CC_w-C_{pw}$ )	0.58	0.76	0.33	0.30
	Intercept	0.65	0.81		
Fixed effects		Estimate	SE	z-value	p
Intercept		0.99	0.41	2.42	0.015**
Group (P&O)		−0.10	0.41	−0.25	ns
TRIAL CONDITION					
( $C_{pw}-C_w$ )		0.49	0.59	0.83	ns
( $CC_w-C_{pw}$ )		1.60	0.64	2.51	0.012*
GROUP*TRIAL CONDITION					
P&O* $C_{pw}-C_w$		1.40	0.70	2.00	0.046*
P&O* $CC_w-C_{pw}$		−1.83	0.61	−2.96	0.003**
Marginal $R^2$ :		0.13	Conditional $R^2$ :	0.41	

\* $p < 0.05$ ; \*\* $p < 0.01$ .

39% more often than with a single consonant (Estimate = 0.33,  $z = 2.33$ ,  $p = 0.020$ ).

## Further Analyses

The fact that the Phonology group spelled some pseudowords with double consonants and some with a singleton consonant led us to hypothesize that this group learned some pseudowords as containing long consonants and others as containing short consonants. If this is true, then these participants should produce longer consonants in pseudowords they spelled with double letters than in pseudowords they spelled with singleton consonant, and consider the former as geminates and the latter as singleton consonants. To test this hypothesis, we categorized each  $CC_{pw}-C_w$  and  $C_{pw}-C_w$  pair produced by each Phonology group participant according to how s/he had spelled the pair in the Spelling Task (see e.g. Sokolović-Perović et al., 2019). If the participant spelled the pseudoword with double letters, we classified the pair as a CC-C pair, and if the participant spelled the pseudoword with a single consonant we classified the pair as a C-C pair. Therefore, we predicted that in both production and awareness the Phonology group's CC-C pairs should behave similarly to the Phonology & Orthography's group  $CC_{pw}-C_w$  pairs, and the former's C-C pairs should behave similarly to the latter's  $C_{pw}-C_w$  pairs.

In the Picture Naming Task, the Phonology group's geometric mean ratios were 1.08 for C-C pairs ( $CI$  [1.02–1.14]) and 1.56 for CC-C pairs ( $CI$  [1.47–1.65]). These ratios were very similar to the Phonology & Orthography group's  $C_{pw}-C_w$  pairs and  $CC_{pw}-C_w$  pairs. We then ran a linear mixed model, inserting Group (Phonology, Phonology & Orthography), Type of Pair (C-C, CC-C), and their interaction as fixed effects,

**TABLE 4 |** Results of mixed-model analysis of the effects of group (Phonology, Phonology & Orthography) on number of pseudowords correctly spelled.

Random effects		Variance	SD		
Participants	Intercept	5.01	2.24		
	Word	1.89	1.38		
Fixed effects		Estimate	SE	z-value	p
Intercept		3.42	0.81	4.24	<0.001***
Group (P&O)		2.65	0.97	2.73	0.006**
Marginal $R^2$ :		0.00	Conditional $R^2$ :	0.06	

\*\* $p < 0.01$ ; \*\*\* $p < 0.001$ .

and the intercepts for subjects and items as random effects. Both Group and the interaction were removed from the model because they did not improve model fit. The effect of Type of Pair (Estimate = 0.35,  $SE = 0.03$ ,  $t = 11.55$ ,  $p < 0.001$ ) showed that C-C pairs were produced with smaller ratios than CC-C for both groups, confirming that participants produced longer consonants when they spelled the pseudoword with double letters.

In the Rhyme Judgment Task, the Phonology group's mean accuracy was higher (81%) for C-C pairs ( $SD = 40$ ,  $CI = 9$ ) than for CC-C pairs ( $M = 66\%$ ,  $SD = 47$ ,  $CI = 9$ ). These figures are very similar to the Phonology & Orthography group's accuracy levels for  $C_{pw}-C_w$  and  $CC_{pw}-C_w$  pairs, respectively. We then ran a generalized linear mixed model with Group (Phonology, Phonology & Orthography), Type of Pair (C-C, CC-C), and their interaction as fixed effects, and the intercepts for subjects and items as random effects. Both Group and the interaction were removed from the model because they did not improve model fit. The effect of Type of Pair (Estimate = −1.24,  $SE = 0.48$ ,  $z = -2.60$ ,  $p = 0.009$ ) shows that both groups were more accurate with C-C than with CC-C pairs.

## DISCUSSION

Orthographic forms affect the perception and production of spoken words in second language learners, but it is not clear whether these effects can be established at the point of first learning the word. The main goal of the current study was then to investigate how the simultaneous presentation of orthographic and phonological inputs affects the learning of novel English<sub>L2</sub> words, and, in particular, we tested whether the presence of a double consonant in the spelling of a new English<sub>L2</sub> word may lead Italian<sub>L1</sub> speakers to perceive and subsequently produce this word with a longer consonant than a word spelled with a singleton consonant, therefore producing a contrast that does not exist in the L2 auditory input or in the L2 phonological system. Results revealed that this was indeed the case.

The first aim of the study was to test whether learning a new word that is spelled with double consonant letters results in producing the new word with a longer consonant than a similar word that is spelled with a singleton consonant letter.

The Phonology & Orthography group, who had learned the novel words' spoken and written form, produced longer consonants in such words, compared to the Phonology group who had learned only the spoken form. This was shown by the interaction between group and type of word pair, and by multiple comparisons between groups.  $CC_{pw}-C_w$  and  $CC_w-C_{pw}$  pairs had a higher consonant duration ratio in the Phonology & Orthography group than in the Phonology group. In addition, the Phonology & Orthography group produced novel CC-words with longer consonants and novel C-words with shorter consonants. This was shown by the high ratios of  $CC_{pw}-C_w$  and  $CC_w-C_{pw}$  pairs (respectively 1.64 and 1.47), while the ratio in  $C_{pw}-C_w$  pairs was just above one (1.13). The high consonant ratios in pairs containing a double consonant show that the double consonant was produced as a geminate, in both real words and newly-learned pseudowords. This confirms previous findings that Italian<sub>L1</sub> speakers of English<sub>L2</sub> produce known English words with a geminate consonant when that consonant is spelled with double letters (Bassetti, 2017; Bassetti et al., 2018), and crucially shows that such effects are found in newly-learned words. The presence of orthographic input during word learning presumably led to the activation of L1 phoneme-grapheme correspondence rules and their transposition to the newly learned L2 words, confirming a strong influence of orthography on L2 production when L1-L2 incongruent graphemes are presented (Hayes-Harb et al., 2010; Pytlyk, 2011; Escudero et al., 2014; Showalter and Hayes-Harb, 2015; Mathieu, 2016).

The Phonology group, who had not seen the items' orthographic forms, produced the novel words' target consonant with similar duration across conditions, and indeed multiple contrasts within group revealed that the mean ratios did not differ across types of pairs. However, crucially, in comparison with the Phonology & Orthography group, the Phonology group produced consonants with a smaller duration ratio in  $CC_{pw}-C_w$  (1.35 vs. 1.64) and  $CC_w-C_{pw}$  pairs (1.14 vs. 1.47), and with a higher ratio in  $C_{pw}-C_w$  pairs (1.30 vs. 1.13). These results support Bassetti's (2017) hypothesis that, in the absence of orthographic input, Italians categorize the duration of English consonants in native speaker's production as in-between short and long consonants, as follows. The Phonology group had a smaller ratio than the Phonology & Orthography group in  $CC_{pw}-C_w$  pairs, because both groups produced C-words with a singleton consonant, but the Phonology & Orthography group generally produced CC-pseudowords with a geminate, and the Phonology group generally produced about half CC-pseudowords with a geminate and half with a singleton. Similarly,  $CC_w-C_{pw}$  pairs had smaller ratios in the Phonology than in the Phonology & Orthography group because both groups produced real CC-words with a geminate, but the Phonology & Orthography group produced all C-pseudowords with a singleton, whereas the Phonology group produced many of these as a geminate. Finally, with  $C_{pw}-C_w$  pairs the Phonology & Orthography group had a smaller ratio than the Phonology group because the former produced all C-pseudowords with a singleton, whereas the latter produced about half of the C-pseudowords with a geminate. Further evidence comes from the re-analysis of the data from the picture naming task, where the Phonology group's stimulus

pairs were categorized according to the participant's spelling as CC-C and C-C pairs, rather than based on the pseudoword's spelling we had taught to the Phonology & Orthography group. When we compared the Phonology group's ratios with pairs they had spelled as CC-C and C-C, they performed similarly to the Phonology & Orthography group's performance with  $CC_{pw}-C_w$  and  $C_{pw}-C_w$  pairs. This finding demonstrated that, in the absence of orthographic input, Italian speakers can perceive English consonants as either singleton or geminate consonants, and this was reflected in both their spoken and written production.

The second aim of the study was to test whether the same orthographic effect on new word learning found in speech production would also be found in metalinguistic awareness. Indeed, the results of the rhyme judgment task were in line with the results from the production task. As predicted, the Phonology & Orthography group incorrectly rejected  $CC_{pw}-C_w$  rhymes more often than  $C_{pw}-C_w$  rhymes, and as often as  $CC_w-C_{pw}$  rhymes, showing that they correctly accepted rhymes where the target consonant is spelled with a singleton letter ( $C_{pw}-C_w$  rhymes, average accuracy of 93%) and incorrectly rejected rhymes where the target consonant is spelled with a singleton letter in one item and double letters in the other one ( $CC_{pw}-C_w$  and  $CC_w-C_{pw}$  rhymes, 69% and 64% correct, respectively). This is presumably because the Phonology & Orthography group erroneously interpreted the presence of a double consonant in word and pseudoword spellings as a long sound and rejected rhymes containing the same consonant spelled with singleton or double consonant, because long and short consonants are different phonemes in their native language. This confirms findings by Bassetti et al. (under revision) with real words, and extends such findings to newly-learned words.

The Phonology group were instead most accurate with  $CC_w-C_{pw}$  rhymes, but showed no difference in accuracy between  $CC_{pw}-C_w$  and  $C_{pw}-C_w$  rhymes. This is because this group—as shown in the production task—interpreted more than half of novel words as containing a geminate consonant. Therefore,  $CC_{pw}-C_w$  and  $C_{pw}-C_w$  rhymes both contained a real C-word and a novel word that was sometimes evaluated to contain a geminate and sometimes a singleton.  $CC_w-C_{pw}$  rhymes were most often accepted because to this group all CC-words contained a geminate and more than half of the pseudowords also contained a geminate. We performed further analysis of the Phonology group's rhyme judgments based on how each participant spelled the pseudowords in the spelling task, as we had done with the data from picture naming, and found that for pairs containing a single-consonant real word the Phonology group rejected more pairs when they had spelled the pseudoword with double consonant letters than when they had spelled the pseudoword with a singleton consonant. This is similar to what the Phonology & Orthography group did with  $CC_{pw}-C_w$  vs.  $C_{pw}-C_w$  rhymes.

It can be concluded that orthographic input affects L2 word acquisition, not only in speech production, but also in a similar way in metalinguistics awareness. This is an interesting finding because there has been very little research on orthographic effects on metalinguistic awareness in second language speakers.



As a third aim, we tested whether the presence of orthographic forms during learning would result in learning more L2 words, compared with auditory input only. Results from the learning and spelling tests show that the Phonology & Orthography group learned more novel words (pseudowords) than the Phonology group. Perhaps more predictably, those who had seen the words' written forms could spell more words than those who had only had auditory input and were guessing the words' spelling. Crucially, the Phonology & Orthography group also learned more spoken words than the Phonology group. This was evident at all four time points during the learning phase. The Phonology & Orthography group learned on average more than 90% of the novel words after just one exposure—compared with about 85% in the Phonology group – and reached ceiling level after just four exposures—compared with around 90% accuracy in the Phonology group. While both groups showed progression in learning, with more words correctly produced in the fourth than in the first repetition, the Phonology group's ultimate performance was similar to the Phonology & Orthography group's performance after first exposure. While previous research found that spoken L2 vocabulary acquisition is more efficient with than without orthographic input in child beginner learners of English (Hu, 2008; Vadasy and Sanders, 2015), the present study shows this facilitative effect in adult experienced L2 learners.

The results of the learning test cannot be considered conclusive, because the analysis is based on self-reported data, whereby participants produced the word from memory, heard the native speaker's model, and evaluated whether they had produced the correct form or not. Future research could use more objective measures of word learning. However, overall our results show that orthographic input results in almost 100% accuracy in spoken and written word learning after just four exposures.

## CONCLUSIONS

This study showed for the first time that the presence of orthographic input during the initial learning of a second language spoken word can lead experienced L2 speakers to learn the phonological form of the word with a sound that does not exist in the auditory input they were exposed to, or indeed in the target language. It is possible that orthographic input results in a perceptual illusion, such that L2 speakers perceive—and therefore learn—a novel word as containing a long consonant if it is spelled with double letters. This is due to recoding the L2 orthographic word using L1 grapheme-phoneme conversion (what Hayes-Harb and others call “orthographic incongruency”). It appears that the effects of orthographic forms on L2 word production that have been widely reported are established in the very early stages of L2 word learning.

While it is perhaps to be expected that those without previous experience of a language's phonology and orthography would fall back on L1 phonology and orthography to make

sense of L2 input, we found such effects in learners with over 10 years' exposure to the second language. Furthermore, orthographic effects have been shown in experienced L2 speakers learning novel words, but these effects resulted in producing or perceiving an incorrect L2 phonological category, whereas here the orthographic effect resulted in the production of a sound that does not exist in the target language. Finally, previous research on orthographic effects on L2 speech production could not explain whether the effect of double letters on consonant length in the production of L2 words was due to repeated exposure or whether it was established at the point of first learning the word. Results from the present study indicated that just four presentations were sufficient to establish a phonological representation containing a sound not present in the auditory input.

We found effects of the same orthographic form (double letters) on speech production, metalinguistic awareness and spelling of newly learned words. These findings confirm Bassetti et al.'s (under revision) findings with real words, and support Bassetti's (2006, 2008) view that L2 phonological representations are affected by both L2 orthographic input, reinterpreted according to L1 orthography-phonology correspondence, and phonological input, reinterpreted according to the L1 phonological system. In this case, learners apply the L1 correspondence between double letters and long consonants, as well as the L1 distinction between long and short consonants, to the English language, where the distinction and correspondence do not exist. The presence of the same orthographic effect on spoken production, written production and metalinguistic awareness is powerful evidence that orthography affects L2 phonological representations, even in newly learned words.

## ETHICS STATEMENT

The study was carried out in accordance with the recommendations of the British Psychological Society. Prior to data collection ethical approval was obtained from the Humanities & Social Sciences Research Ethics Committee of the University of Warwick (UK). All participants signed a written informed consent form.

## AUTHOR CONTRIBUTIONS

TC and BB contributed to the design of the study and the interpretation of results. TC was responsible for data collection and analysis and drafted the manuscript. BB and JM revised it critically for important intellectual content. All the authors agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

## FUNDING

This study was supported by a Leverhulme Trust Research Grant [grant RPG 2013 180] awarded to BB and JM.

## ACKNOWLEDGMENTS

The authors are grateful to Paolo Mairano and Rosalba Nodari for performing the acoustic analysis, to Jo Gakonga for recording auditory stimuli, and to all participants and their teachers.

## REFERENCES

- Baayen, R. H., Davidson, D. J., and Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *J. Memory Lang.* 59, 390–412. doi: 10.1016/j.jml.2007.12.005
- Barr, D. J., Levy, R., Scheepers, C., and Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: keep it maximal. *J. Memory Lang.* 68, 255–278. doi: 10.1016/j.jml.2012.11.001
- Bartoň, K. (2018). *MuMIn: Multi-Model Inference. Package version 1.42.1*. Available online at: <https://CRAN.R-project.org/package=MumIn>
- Basetti, B. (2006). Orthographic input and phonological representations in learners of Chinese as a foreign language. *Written Lang. Liter.* 9, 95–114. doi: 10.1075/wll.9.1.07bas
- Basetti, B. (2007). “Effects of hanyu pinyin on pronunciation in learners of Chinese as a Foreign Language,” in *The Cognition, Learning and Teaching of Chinese Characters*, eds A. Guder, X. Jiang and Y. Wan (Beijing: Beijing Language and Culture University Press), 156–179.
- Basetti, B. (2008). “Orthographic input and second language phonology,” in *Input Matters in SLA* eds T. Piske and M. Young-Scholten (Clevedon: Multilingual Matters), 191–206. doi: 10.21832/9781847691118-013
- Basetti, B. (2017). Orthography affects second language speech: double letters and geminate production in English. *J. Exp. Psychol. Learn. Memory Cogn.* 43, 1835–1842. doi: 10.1037/xlm0000417
- Basetti, B., and Atkinson, N. (2015). Effects of orthographic forms on pronunciation in experienced instructed second language learners. *Appl. Psycholinguistics* 36, 67–91. doi: 10.1017/S0142716414000435
- Basetti, B., Sokolović-Perović, M., Mairano, P., and Cerni, T. (2018). Orthography-induced length contrasts in the second language phonological systems of L2 speakers of English: Evidence from minimal pairs. *Lang. Speech* 61, 577–597. doi: 10.1177/0023830918780141
- Bates, D., Kliegl, R., Vasishth, S., and Baayen, H. (2015a). *Parsimonious Mixed Models*. arXiv preprint arXiv:1506.04967.
- Bates, D., Maechler, M., Bolker, B., and Walker, S. (2015b). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48. doi: 10.18637/jss.v067.i01
- Boersma, P., and Weenink, D. (2016). *Praat: Doing Phonetics by Computer [Computer program]*. Version 6.0.19. Available online at: <http://www.praat.org/>
- Chambré, S. J., Ehri, L. C., and Ness, M. (2017). Orthographic facilitation of first graders’ vocabulary learning: does directing attention to print enhance the effect? *Read. Writ.* 30, 1137–1156. doi: 10.1007/s11145-016-9715-z
- Detey, S., and Nespoulous, J. (2008). Can orthography influence second language syllabic segmentation? Japanese epenthetic vowels and French consonantal clusters. *Lingua* 118, 66–81. doi: 10.1016/j.lingua.2007.04.003
- Ehri, L. C., and Wilce, L. S. (1979). The mnemonic value of orthography among beginning readers. *J. Educ. Psychol.* 71, 26–40. doi: 10.1037/0022-0663.71.1.26
- Erdener, V. D., and Burnham, D. K. (2005). The role of audiovisual speech and orthographic information in nonnative speech production. *Lang. Learn.* 55, 191–228. doi: 10.1111/j.0023-8333.2005.00303.x
- Escudero, P., Hayes-Harb, R., and Mitterer, H. (2008). Novel second-language words and asymmetric lexical access. *J. Phonet.* 36, 345–360. doi: 10.1016/j.wocn.2007.11.002
- Escudero, P., Simon, E., and Mulak, K. E. (2014). Learning words in a new language: orthography doesn’t always help. *Bilingualism Lang. Cogn.* 17, 384–395. doi: 10.1017/S1366728913000436
- Flege, J. E. (1995). “Second language speech learning: theory, findings and problems,” in *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues*, ed W. Strange (Timonium, MD: York Press), 233–277.
- Hayes-Harb, R., Nicol, J., and Barker, J. (2010). Learning the phonological forms of new words: effects of orthographic and auditory input. *Lang. Speech* 53, 367–381. doi: 10.1177/0023830910371460
- Hu, C.-F. (2008). Use orthography in L2 auditory word learning: who benefits? *Read. Writing* 21, 823–841. doi: 10.1007/s11145-007-9094-6
- Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B. (2017). lmerTest Package: tests in linear mixed effects models. *J. Stat. Soft.* 82, 1–26. doi: 10.18637/jss.v082.i13
- Lenth, R. V. (2016). Least-squares means: the R package lsmeans. *J. Stat. Softw.* 69, 1–33. doi: 10.18637/jss.v069.i01
- Mathôt, S., Schreij, D., and Theeuwes, J. (2012). OpenSesame: an open-source, graphical experiment builder for the social sciences. *Behav. Res. Methods* 44, 314–324. doi: 10.3758/s13428-011-0168-7.
- Mathieu, L. (2016). The influence of foreign scripts on the acquisition of a second language phonological contrast. *Second Lang. Res.* 32, 145–170. doi: 10.1177/0267658315601882
- Miles, K. P., Ehri, L. C., and Lauterbach, M. D. (2016). Mnemonic value of orthography for vocabulary learning in monolinguals and language minority English-speaking college students. *J. College Read. Learn.* 46, 99–112. doi: 10.1080/10790195.2015.1125818
- Nelson, J. R., Balass, M., and Perfetti, C. A. (2005). Differences between written and spoken input in learning new words. *Written Lang. Literacy* 8, 25–44. doi: 10.1075/wll.8.2.04nel
- Nova Development (2004). *Art Explosion Library*. Calabasas, CA: Nova Development Co.
- Pytyk, C. (2011). Shared orthography: do shared written symbols influence the perception of L2 sounds? *Modern Lang. J.* 54, 541–557. doi: 10.2307/41413379
- R Core Team (2018). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Available online at: <http://www.R-project.org/>
- Ricketts, J., Bishop, D. V., and Nation, K. (2009). Orthographic facilitation in oral vocabulary acquisition. *Q. J. Exp. Psychol.* 62, 1948–1966. doi: 10.1080/17470210802696104
- Rosenthal, J., and Ehri, L. C. (2008). The mnemonic value of orthography for vocabulary learning. *J. Educ. Psychol.* 100:175. doi: 10.1016/j.jecp.2015.01.015
- RStudio Team (2018). *RStudio: Integrated Development for R*. Boston, MA: RStudio, Inc. Available online at: <http://www.rstudio.com/>
- Showalter, C., E., and Hayes-Harb, R. (2013). Unfamiliar orthographic information and second language word learning: a novel lexicon study. *Second Lang. Res.* 29, 185–200. doi: 10.1177/0267658313480154
- Showalter, C. E., and Hayes-Harb, R. (2015). Native English speakers learning Arabic: the influence of novel orthographic information on second language phonological acquisition. *Appl. Psycholinguistics* 36, 23–42. doi: 10.1017/S0142716414000411
- Sokolović-Perović, M., Basetti, B., and Dillon, S. (2019). English orthographic forms affect L2 English speech production in native users of a non-alphabetic writing system. *Bilingual. Lang. Cogn.* 1–11. doi: 10.1017/S136672891900035X
- Vadasy, P. F., and Sanders, E. A. (2015). Incremental learning of difficult words in story contexts: the role of spelling and pronouncing new

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcomm.2019.00031/full#supplementary-material>

- vocabulary. *Read. Writing* 28, 371–394. doi: 10.1007/s11145-014-959-9
- Vokic, G. (2011). When alphabets collide: alphabetic first-language speakers' approach to speech production in an alphabetic second language. *Second Lang. Res.* 27, 391–417. doi: 10.1177/0267658310396627
- Young-Scholten, M., Akita, M., and Cross, N. (1999). "Focus on form in phonology: orthographic exposure as a promoter of epenthesis," in *Pragmatics and pedagogy. Proceedings of the third PacSLRF*, eds P. Robinson and N. O. Jungheim (Tokyo: Aoyama Gakuin University), 2, 227–233.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Cerni, Bassetti and Masterson. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.